AEC Computing and Applied Mathematics Center

AEC RESEARCH AND DEVELOPMENT REPORT

TID-4500
14th Ed.

NYO-2878
PHYSICS

A DUALITY THEOREM FOR CONVEX PROGRAMS

by

W. S. Dorn

August 1, 1959

NYU NYO-2878 c.3
Dorn
A duality theorem for
convex programs.

NEW YORK UNIVERSITY
COURANT INSTITUTE - LIBRARY
251 Mercer St. New York, N.Y. 10012

Institute of Mathematical Sciences

NEW YORK UNIVERSITY
NEW YORK, NEW YORK

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, express or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

AEC Computing and Applied Mathematics Center
Institute of Mathematical Sciences
New York University

TID-4500
14th Ed.

NYO-2878
PHYSICS

A DUALITY THEOREM FOR CONVEX PROGRAMS

by

W. S. Dorn

August 1, 1959

Contract No. AT(30-1)-1480

NYO-2878
PHYSICS AND
MATHEMATICS

ABSTRACT

A proof is given for a dual theorem for a class of convex programs, i.e. constrained minimization of convex functions. A simple example is included.

TABLE OF CONTENTS

Page

Abstract 2

1. Introduction 4
2. Notation 4
3. Convex Functions 5
4. Duality in Linear Programming 6
5. A Class of Convex Programs and Its Dual . . . 7
6. A Duality Theorem for Convex Programs 8
7. Remarks 16

Bibliography 22

A DUALITY THEOREM FOR CONVEX PROGRAMS

1. Introduction.

In a recent paper [1], the duality relationship of linear programming was extended to a class of quadratic programs. In what follows this concept will be extended further to include convex programs, i.e., programs to minimize a convex function subject to linear constraints. The arguments, to some extent, parallel those given in the quadratic case. The notion of duality for this class of programs has also been discussed previously by Dennis [2].

2. Notation.

In what follows matrix notation will be used. Lower case letters $x,y,\ldots$ will denote column vectors and capital letters $A,C,\ldots$ will be matrices. The operator ∇ is the gradient operator and $\nabla f(x)$ is a vector composed of the first partial derivatives of the scalar function $f(x)$.

Prime denotes transpose so that $x',y',\ldots$ are row vectors. The product $x'y$ is the inner product of the vectors x and y. A vector inequality will apply to each component of the vector, i.e. $x \geq 0$ implies that each component of x is non-negative.

3. Convex Functions.

A scalar function, f, of an n-dimensional vector, x, is convex if linear interpolation does not underestimate the value of the function, i.e.,

$$f(x_1 + k[x_2-x_1]) \leq f(x_1) + k[f(x_2) - f(x_1)]$$

for $0 \leq k \leq 1$. For convex functions a local minimum is also a global minimum.

For the proofs in the succeeding sections the following lemma[1] regarding convex functions is required.

Lemma. If $f(x)$ is convex then for any vectors y and x

$$f(y) - f(x) \geq (y-x)' \nabla f(x) \tag{3.1}$$

Proof. From the assumption of convexity

$$f(x + k[y-x]) \leq f(x) + k[f(y) - f(x)]$$

for all k such that $0 \leq k \leq 1$. Rearranging terms

$$k[f(y) - f(x)] \geq f(x + k[y-x]) - f(x) . \tag{3.2}$$

From the differential mean value theorem

$$f(x + k[y-x]) - f(x) = k(y-x)' \nabla f(x+x_o)$$

where $x_o = \theta k(y-x)$ for some θ such that $0 \leq \theta \leq 1$.

---

[1] The author is indebted to C. E. Lemke for this lemma.

Using this in the right hand member of (3.2)

$$f(y) - f(x) \geq (y-x)' \nabla f(x + x_o)$$

Letting k approach zero, the inequality (3.1) results.

4. Duality in Linear Programming.

The linear programming problem may be posed as follows: To minimize the linear form $p'x$ over all n-dimensional vectors x satisfying the constraints

$$A x \geq b \qquad (4.1)$$

$$x \geq 0 \qquad (4.2)$$

where p is an $n \times 1$ vector, b is an $m \times 1$ vector A is an $m \times n$ matrix.

The dual problem to the above is to maximize $b'v$ over all m-dimensional vectors v satisfying

$$A'v \leq p \qquad (4.3)$$

$$v \geq 0 \qquad (4.4)$$

The duality theorem [3] states that if a solution to either problem exists and is finite, then a solution to the other problem also exists and indeed

$$\text{Minimum } p'x = \text{Maximum } b'v \qquad (4.5)$$

The constraints (4.1), (4.2) in the primal problem may take on other forms with corresponding changes in the dual constraints (4.3), (4.4). One form which will be of particular interest is given here. Let the vector x be partitioned as (x_1,x_2) and consider the problem of minimizing $p_1'x_1 + p_2'x_2$ subject to

$$A_1x_1 + A_2x_2 \geq b \qquad (4.6)$$

$$x_1 \geq 0 \qquad (4.7)$$

where p and A have been similarly partitioned. Note that x_2 is not required to be non-negative. These constraints can be put into the form of (4.1) and (4.2) by the substitution

$$x_2 = z - y$$

where $z \geq 0$, $y \geq 0$. All variables x_1 , z and y are now non-negative.

The dual problem is then from (4.3) and (4.4)

$$\text{Maximize} \quad b'v \qquad (4.8)$$

subject to

$$A_1' v \leq p_1 \qquad (4.9)$$

$$A_2' v = p_2 \qquad (4.10)$$

$$v \geq 0 \qquad (4.11)$$

5. A Class of Convex Programs and Its Dual.

Consider the following class of non-linear programming problems:

$$\text{Minimize:} \quad \phi(x) = f(x) + p'x \qquad (5.1)$$

subject to

$$Ax \geq b \tag{5.2}$$

$$x \geq 0 \tag{5.3}$$

where $f(x)$ is a convex scalar function of the vector x and possesses continuous first partial derivatives with respect to each component of x. Moreover, it will be assumed that given a vector y, the vector equation $\nabla f(x) = y$ possesses a solution x which is one differentiable with respect to y. The vectors p, b, and x and the matrix A are the same as in Section 3. Denote this as Problem I.

A dual problem to Problem I is

$$\text{Maximize} \quad g(u,v) = f(u) - u' \nabla f(u) + b'v \tag{5.4}$$

subject to

$$A'v - \nabla f(u) \leq p \tag{5.5}$$

$$v \geq 0 \tag{5.6}$$

Denote this as Problem II.

Problems I and II are related through the theorem in the next section.

6. A Duality Theorem for Convex Programs.

Theorem: (i) If there exists a vector x_o which minimizes $\phi(x)$ in Problem I, then there also exists vectors $u = x_o$, $v = v_o$ which maximize $g(u,v)$ in Problem II.

(ii) Conversely, if u_o, v_o are vectors which maximize $g(u,v)$ in Problem II, then $x = u_o$ minimizes $\phi(x)$ in Problem I. In either case

$$\text{Minimum } \phi(x) = \text{Maximum } g(u,v) .$$

Proof: (i) Let x_o be a solution to Problem I, then

$$\phi(x_o) \leq \phi(x) \tag{6.1}$$

Consider the following linear programming problem

$$\text{Minimize } F(x) = -f(x_o) + x' \nabla f(x_o) + x'p \tag{6.2}$$

$$Ax \geq b \tag{6.3}$$

$$x \geq 0 \tag{6.4}$$

Call this Problem I'. Note that x_o is a feasible solution to Problem I' in the sense that it satisfies the constraints (6.3) and (6.4). Suppose there exists another feasible solution x^* to Problem I' with the property that

$$F(x^*) < F(x_o) ,$$

i.e.,

$$(x^*-x_o)' [\nabla f(x_o)+p] < 0 \tag{6.5}$$

It is easily verified that

$$x_1 = x_o + k(x^*-x_o) , \qquad 0 \leq k \leq 1$$

is also a feasible solution to both Problem I and Problem I'. Consider

$$\phi(x_1) - \phi(x_o) = [f(x_1)+p'x_1] - [f(x_o)+p'x_o]$$

From the mean value theorem

$$\phi(x_1) - \phi(x_o) = (x_1-x_o)' [\nabla f(x_o-\theta\{x_1-x_o\})+p] , \quad 0\leq\theta\leq 1$$

$$= k(x^*-x_o)' [\nabla f(x_o-\theta k\{x^*-x_o\})+p]$$

$$= k\Big\{(x^*-x_o)'[\nabla f(x_o-\theta k[x^*-x_o])-\nabla f(x_o)]$$

$$+ (x^*-x_o)' [\nabla f(x_o) + p]\Big\}$$

The second term in the braces is negative by virute of the assumption (6.5). Moreover, this term is independent of k. However, since $\nabla f(x)$ is continuous, for k sufficiently small the first term in the braces may be made arbitrarily small. In particular k may be chosen so that

$$\left|(x^*-x_o)'[\nabla f(x_o-\theta k[x^*-x_o])-\nabla f(x_o)]\right| < \left|(x^*-x_o)'[\nabla f(x_o)+p]\right|$$

It follows then from the inequality (6.5) that

$$\phi(x_1) - \phi(x_o) < 0$$

But this is in contradiction to(6.1). Therefore, x_o must be the minimizing solution of Problem I'.

From the duality theorem of linear programming a solution also exists to the dual to Problem I'. This dual linear problem is from (4.3) and (4.4).

$$\text{Maximize} \quad G(v) = - f(x_o) + b'v \qquad (6.6)$$

$$A'v \leq f(x_o) + p \qquad (6.7)$$

$$v \geq 0 \qquad (6.8)$$

Call this Problem II'. Denote the vector v which maximizes Problem II' by v_o. Then the duality theorem of linear programming, (4.5), requires that

$$x_o' \nabla f(x_o) + x_o'p = b'v_o . \qquad (6.9)$$

Notice now that $u = x_o$, $v = v_o$ comprises a feasible solution to Problem II. To show that it also provides a maximal solution consider

$$g(x_o,v_o) - g(u,v) = f(x_o)-x_o'\nabla f(x_o)+b'v_o-f(u)+u'\nabla f(u)-b'v$$

From (6.9)

$$g(x_o,v_o) - g(u,v) = f(x_o) - f(u)+x_o'p+u'\nabla f(u)-b'v$$

and from the lemma

$$g(x_o,v_o)-g(u,v) \geq (x_o-u)'\nabla f(u)+x_o'p+u'\nabla f(u)-b'v$$

$$= x_o' [\nabla f(u) + p] - b'v \qquad (6.10)$$

Now from (5.2) and (5.6)

$$- b'v \geq -x_o' A' v$$

and from (5.3) and (5.5)

$$x_o' [\nabla f(u) + p] \geq x_o' A'v$$

Substituting these last two inequalities in (6.10)

$$g(x_o,v_o) - g(u,v) \geq x_o' A'v - x_o' A'v = 0$$

Thus (x_o,v_o) is a maximizing solution for Problem II.

Finally from (5.4), (6.9) and (5.1)

$$g(x_o,v_o) = f(x_o)-x_o'\nabla f(x_o)+b'v_o = f(x_o)+x_o'p = \phi(x_o)$$

which verifies the equality of the objective functions.

(ii) Let (u_o, v_o, y_o) be a maximizing solution of Problem II rephrased as

$$\text{Maximize} \quad \psi(u,v,y) = f(u) - u'y + b'v \tag{6.11}$$

$$A'v - y \leq p \tag{6.12}$$

$$y = \nabla f(u) \tag{6.13}$$

$$v \geq 0 \tag{6.14}$$

Thus

$$\psi(u_o, v_o, y_o) \geq \psi(u,v,y) \tag{6.15}$$

for all u, v, y satisfying the constraints.

Consider the related linear problem

$$\text{Maximize} \quad G(v,y) = f(u_o) - u_o'y + b'v$$

$$A'v - y \leq p$$

$$v \geq 0$$

Denote this as Problem II*.

Notice that $v = v_o$, $y = y_o$ is a feasible solution to Problem II*. Suppose there exists v^*, y^* satisfying (6.12) and (6.14) such that

$$G(v^*, y^*) > G(v_o, y_o) ,$$

i.e.,

$$-u_o'(y^* - y_o) + b'(v^* - v_o) > 0 \tag{6.16}$$

Define vectors v_1, y_1 to be

$$v_1 = v_o + k(v^* - v_o) \tag{6.17}$$

$$y_1 = y_o + k(y^* - y_o) \tag{6.18}$$

for $0 < k \leq 1$. Then v_1, y_1 is also a feasible solution to Problem II*. Now by assumption (Section 5) there exists a vector u_1 such that

$$y_1 = \nabla f(u_1) \tag{6.19}$$

and thus u_1, v_1, y_1 is a feasible solution to Problem II. Consider

$$\psi(u_1,v_1,y_1) - \psi(u_o,v_o,y_o) = f(u_1) - u_1'y_1 + b'v_1 - f(u_o) + u_o'y_o - b'v_o$$

From the mean value theorem

$$\psi(u_1,v_1,y_1) - \psi(u_o,v_o,y_o) = (u_1-u_o)' \nabla f(u_1-\theta[u_1-u_o]) - u_1'y_1 + u_o'y_o + b'(v_1-v_o) \quad 0 \leq \theta \leq 1$$

and using equation (6.19)

$$\psi(u_1,v_1,y_1) - \psi(u_o,v_o,y_o) = (u_1-u_o)'\left\{\nabla f(u_1-\theta[u_1-u_o]) - \nabla f(u_1)\right\} - u_o'(y_1-y_o) + b'(v_1-v_o) \tag{6.20}$$

Since u possesses first derivatives with respect to y, where $y = \nabla f(u)$, it follows from the mean value theorem that

$$(u_1-u_o)' = (y_1-y_o)' A(y_o - \rho[y_1-y_o]) , \qquad 0 \leq \rho \leq 1$$

where $A(y)$ is a matrix whose i,j element is the partial derivative of the jth component of u with respect to the ith component of y. Using this and equations (6.17) and (6.18), equation (6.20) becomes

$$\psi(u_1,v_1,y_1) - \psi(u_o,v_o,y_o) =$$

$$k[(y^*-y_o)' A(y_o-\rho k[y_1-y_o])\left\{\nabla f(u_1-\theta k[u^*-u_o])-\nabla f(u_1)\right\}$$

$$- u_o'(y^*-y_o) + b'(v^*-v_o)] \tag{6.21}$$

Now if k is sufficiently small, the following inequality can be satisfied.

$$|(y^*-y_o)' A(y_o-\rho k[y^*-y_o])\left\{\nabla f(u_1-\theta k[u^*-u_o]) - \nabla f(u_1)\right\}|$$

$$< |- u_o'(y^*-y_o) + b'(v^*-v_o)|$$

since $\nabla f(u)$ is continuous and the term on the right is independent of k.

From the assumption (6.16), the term inside the absolute value signs on the right of the above inequality is positive. Thus the term in square brackets in equation (6.21) may be made positive by a sufficiently small choice of k. This implies that

$$\psi(u_1,v_1,y_1) - \psi(u_o,v_o,y_o) > 0$$

and contradicts the hypothesis (6.15) that u_o,v_o,y_o is a maximizing solution of Problem II. It follows that v_o,y_o maximizes Problem II*.

Problem II^* is a linear problem of the second type mentioned in Section 4, equations (4.6) and (4.7), and its dual problem according to (4.8) to (4.11) is

$$\text{Minimize} \quad F(x) = f(u_o) + p'x \tag{6.22}$$

$$Ax \geq b \tag{6.23}$$

$$x = u_o \tag{6.24}$$

$$x \geq 0 \tag{6.25}$$

This will be referred to as Problem I^*. The duality theorem of linear programming assures the existence of a solution to this problem. By virtue of (6.24) u_o must be that solution. The duality theorem, (4.5), then provides that

$$\text{Maximum} \quad G(v,y) = \text{Minimum} \quad F(x)$$

or

$$- u_o'y_o + b'v_o = p'u_o \tag{6.26}$$

Now u_o, since it satisfies (6.23) and (6.25), is a feasible solution to Problem I. Let x be any other feasible solution to Problem I and consider

$$\phi(x) - \phi(u_o) = f(x) - f(u_o) + p'(x-u_o)$$

From the lemma again

$$\phi(x) - \phi(u_o) \geq (x-u_o)'[\nabla f(u_o)+p] \tag{6.27}$$

Now from (6.13) and (6.26)

$$- u_o' [\nabla f(u_o)+p] = - u_o'(y_o+p) = - b'v_o$$

and from (5.2) and (5.6)

$$- b'v_o \geq - x' A' v_o$$

so

$$- u_o' [\nabla f(u_o) + p] \geq - x' A' v_o \qquad (6.28)$$

From (5.3) and (5.5)

$$x' [\nabla f(u_o) + p] \geq x' A' v_o \qquad (6.29)$$

Substituting (6.28) and (6.29) into the right hand member of (6.27) it follows that

$$\phi(x) - \phi(u_o) \geq x'A'v_o - x'A'v_o = 0$$

Thus u_o minimizes Problem I.

From (5.1), (6.26) and (5.4)

$$\phi(u_o) = f(u_o)+p'u_o = f(u_o)-u_o' \nabla f(u_o)+b'v_o = g(u_o,v_o)$$

7. Remarks.

It should be noted that the rather strong assumption (Section 5) that $y = \nabla f(x)$ has a differentiable solution x for all y is necessary only for the proof of part (ii) of the theorem. That is to say, the existence of a solution x_o to Problem I implies the existence of a solution x_o, v_o to Problem II even when this assumption is not valid.

The significance of the assumption may be seen from the quadratic case where $f(x) = 1/2 (x'Cx)$. The above assumption then requires that C be positive, definite.

The duality theorem for quadratic programs [1] requires only that C be positive, semi-definite and the duality theorem for convex programs presented here is, in that sense, less general than the corresponding quadratic theorem.

The proof of Section 6 provides a method for computing the dual variables v_o since the primal variables x_o have been obtained. Indeed, v_o is the solution of the following linear programming problem

$$\text{Maximize} \quad b'v$$

subject to

$$A'v \leq p + \nabla f(x_o)$$

$$v \geq 0$$

Finally the constraints of the primal problem (Problem I) may be formulated in other ways with corresponding changes in the dual constraints. Some of these are tabulated below. In all cases the functionals to be optimized are the same as those in Section 5.

Problem I	Type	Problem II
$Ax \geq b$ $x \geq 0$	I	$A'v - \nabla f(u) \leq p$ $v \geq 0$
$Ax \geq b$	II	$A'v - \nabla f(u) = p$ $v \geq 0$
$Ax = b$ $x \geq 0$	III	$A'v - \nabla f(u) \leq p$
$Ax = b$	IV	$A'v - \nabla f(u) = p$

8. Example

As a simple example of the theorem of Section 6, consider the problem of minimizing

$$f(x_1,x_2) = -\log x_1 - \log x_2$$

subject to

$$-x_1 - x_2 \geq -2$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

The admissible values of x_1,x_2 lie in the triangle OAB of Figure I. The contours of constant values of f are also shown with f decreasing as the contours recede from the origin.

The optimal value is $f = 0$ for $x_1 = x_2 = 1$.

The dual problem is

$$\text{Maximize } g(u_1,u_2,v) = -\log u_1 - \log u_2 + 2 - 2v$$

where

$$-v + \frac{1}{u_1} \leq 0$$

$$-v + \frac{1}{u_2} \leq 0$$

$$v \geq 0$$

From the symmetry of the dual it is clear that when g is a maximum, $u_1 = u_2$. The admissible values of u_1 and v are those above the hyperbola CD in Figure 2. The

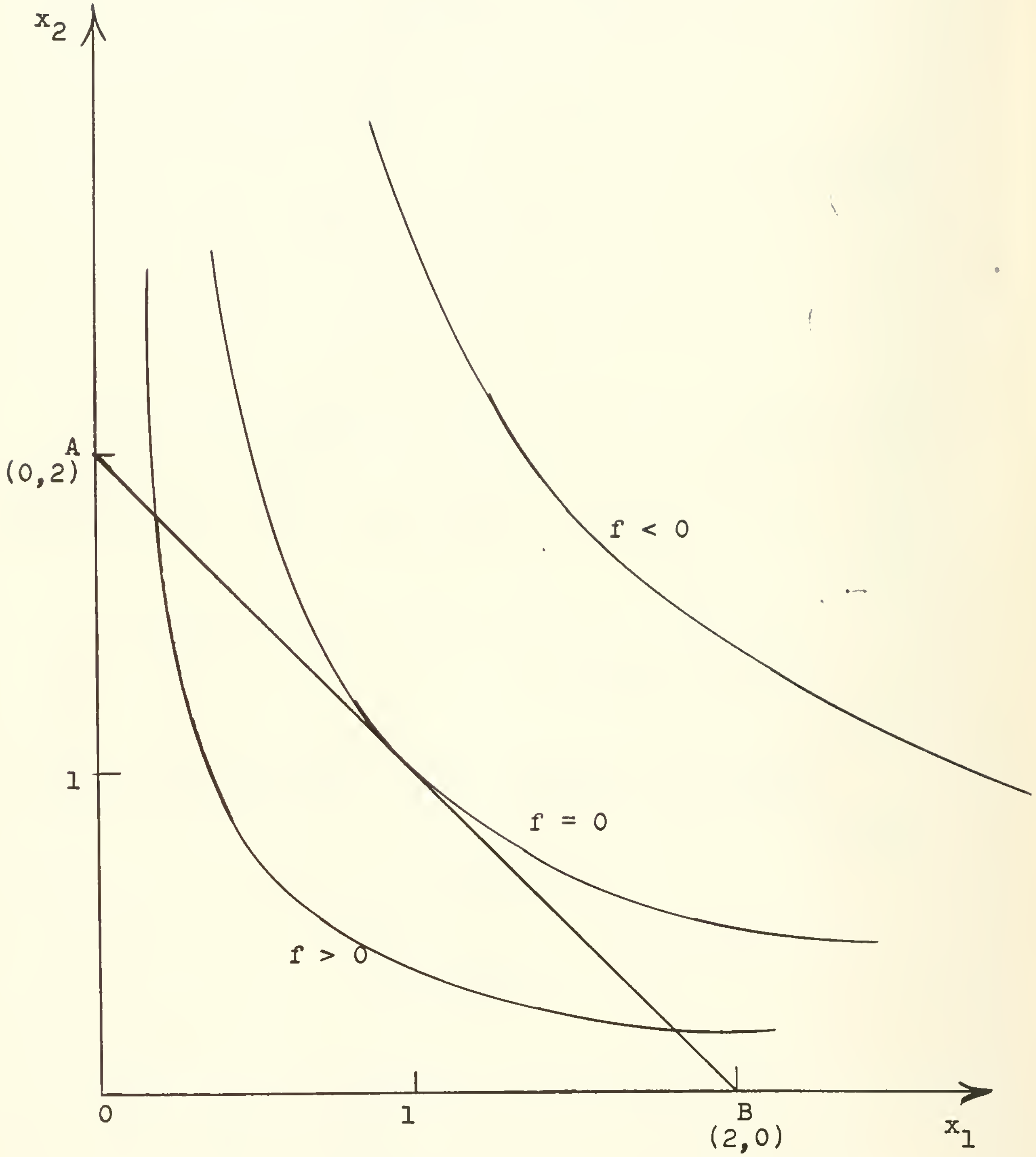

Figure 1

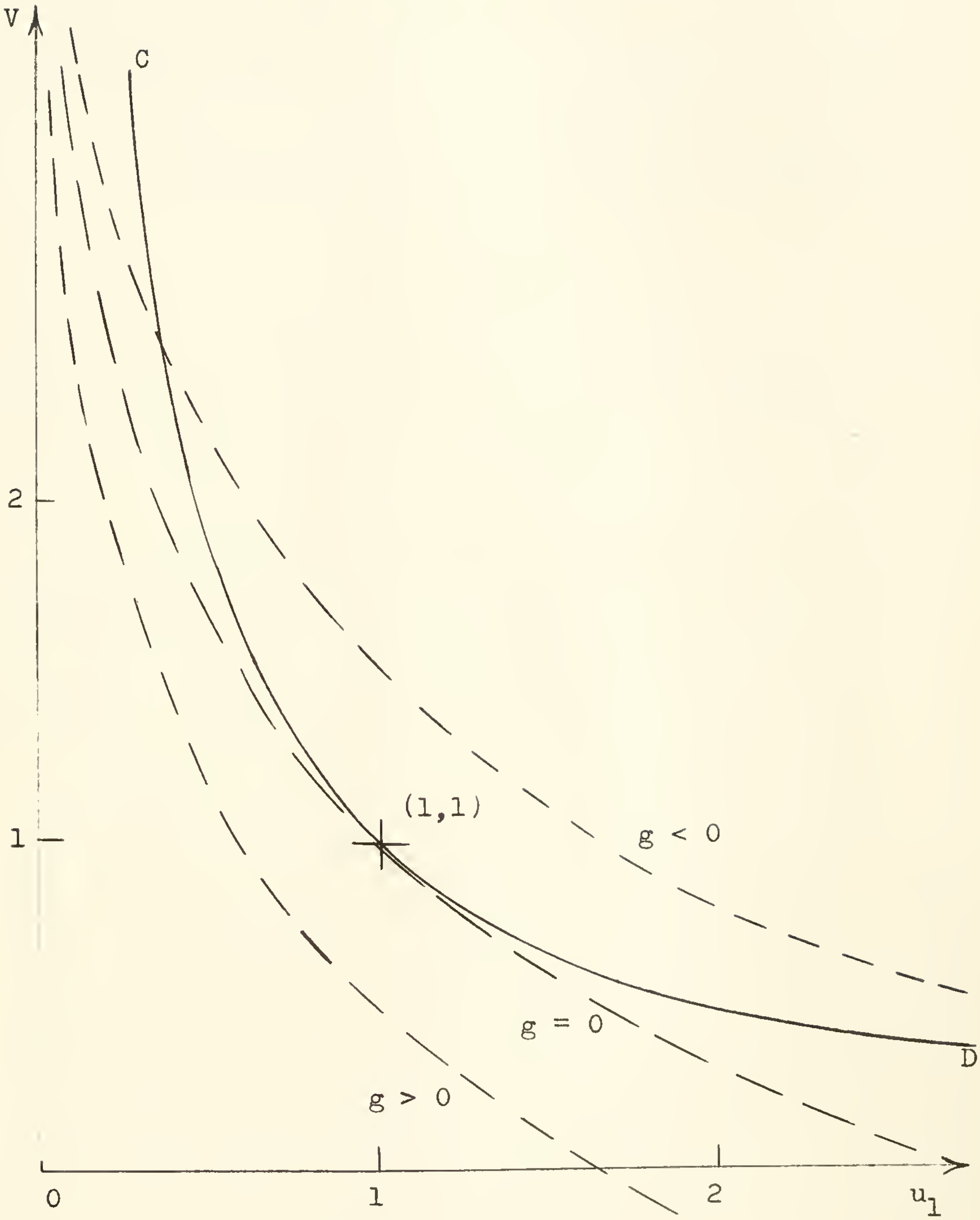

Figure 1

contours of constant g are shown as dotted curves with g increasing as the contours approach the origin. The maximum value of g is associated with the contour which is tangent to the hyperbola CD at $u_1 = u_2 = v = 1$. For these values of the variables $g(1,1,1) = 0$.

Bibliography

1 Dorn, W. S., Duality in Quadratic Programming, AEC Computing and Applied Mathematics Center Report, NYO-8676 (October 1958). To appear in Quarterly of Applied Mathematics.

2 Dennis, Jack B., Mathematical Programming and Electrical Networks, MIT Interim Technical Report No. 10 (September 1958).

3 Gale, D., Kuhn, H. W., Tucker, A. W., Linear Programming and the Theory of Games, Chapter XIX, Activity Anal. of Prod. and Alloc., Cowles Comm. Mono. 13, John Wiley (1951).

NYU NYO-2878 c.3
Dorn
A duality theorem for
convex programs.

NYU NYO-2878 c.3
Dorn
A duality theorem for
convex programs.

This book may be kept
FOURTEEN DAYS
A fine will be charged for each day the book is kept overtime.

NOV 18 [illegible]

GAYLORD 142 PRINTED IN U.S.A.